

RECEIVED

1631

NOV 27 2002

TECH CENTER 1600/2900



1600

#11

RAW SEQUENCE LISTING

DATE: 11/20/2002

PATENT APPLICATION: US/09/724,676A

TIME: 10:48:50

Input Set : D:\pat.ST25.txt

Output Set: N:\CRF4\11192002\I724676A.raw

3 <110> APPLICANT: Compugen LTD
 5 <120> TITLE OF INVENTION: Variants of alternative splicing
 7 <130> FILE REFERENCE: 129181.4 Compugen
 9 <140> CURRENT APPLICATION NUMBER: 09/724,676A
 10 <141> CURRENT FILING DATE: 2000-11-28
 12 <160> NUMBER OF SEQ ID NOS: 97222
 14 <170> SOFTWARE: PatentIn version 3.2

16 <210> SEQ ID NO: 1
 17 <211> LENGTH: 745
 18 <212> TYPE: DNA
 19 <213> ORGANISM: Homo sapiens
 21 <400> SEQUENCE: 1

ENTERED

22	cgggggctgcc	cttcttccct	gcgaggaggag	ggcctgggcg	gtcgcgttgg	cgggaggagg	60
24	gttacctttc	ccagtctcgc	tctggccgcc	tgagccagga	ggaagcagcg	gcgaggtctg	120
26	cgggaggcat	ggcgggagct	ccggacgagc	gccggcgggg	ccccgcggca	ggggagcagc	180
28	tgcagcagca	acacgtctct	tgccaggtct	tccccgagcg	tctggcccag	gggaatcccc	240
30	agcaagggtt	cttctccagc	ttcttcacca	gcaaccagaa	gtgccagctt	aggctcctga	300
32	agacgctgga	gacaagtagg	agccatgacc	tggaggctgt	ggttcacag	aatgggtctg	360
34	agaccgggtg	ggcgaggaag	ggcctgggca	acacctggcc	aggggcttca	ggttcagcac	420
36	agagtctaga	cagactgggc	atcatggggg	ctgggttggg	agcctgaggc	tcagtttccc	480
38	cactagaggt	ggcgtgaga	cccttttcgt	cgtgttttt	ggcttgtgga	atgaatttgt	540
40	ctgccctttt	ttgtcactca	ctgttgtgac	agatcctccg	atttcttaca	taagttatgc	600
42	agccacactg	tgacttcagg	gtaagggtga	ggccaagttt	taaagtgcc	gccagggttaa	660
44	ttaaggcaca	caaaggctgt	aaaaatgtaa	aaatgtaatt	aaggcatacc	aaggctgtaa	720
46	aaatataaaa	gcagttttgc	aaacc				745

49 <210> SEQ ID NO: 2
 50 <211> LENGTH: 754
 51 <212> TYPE: DNA
 52 <213> ORGANISM: Homo sapiens
 54 <220> FEATURE:
 55 <221> NAME/KEY: misc_feature
 56 <222> LOCATION: (14)..(14)
 57 <223> OTHER INFORMATION: n is a,c,g, or t
 59 <220> FEATURE:
 60 <221> NAME/KEY: misc_feature
 61 <222> LOCATION: (23)..(23)
 62 <223> OTHER INFORMATION: n is a,c,g, or t
 64 <220> FEATURE:
 65 <221> NAME/KEY: misc_feature
 66 <222> LOCATION: (25)..(26)
 67 <223> OTHER INFORMATION: n is a,c,g, or t
 69 <400> SEQUENCE: 2

W--> 70 tgcattggaat ttgnacgaga cntnntgaa ccaccggcct gtgactgctg tggagatcca 60

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/724,676A

DATE: 11/20/2002

TIME: 10:48:50

Input Set : D:\pat.ST25.txt

Output Set: N:\CRF4\11192002\I724676A.raw

```

72 gctgatggtg gaagagagtg aagagcggct cacggaggag cagattgaag ctcttctcca 120
74 caccgtcacc agcattctgc ctgcagagcc agaggctgag cagaagaaga atacaaacag 180
76 caatgtggca atggacgaag aggacccagc atagaagagc acagctggcc cggcggttcc 240
78 atgaagtcag aaggcctggc agccatttcc tggacgttga gaggattgtt tatttgattt 300
80 ttatcctcat cccagcaggc ctggctttgt ggtagttgg gtacatcaca aaaataagtt 360
82 aaaaagaaat atttgtgcct tggggagaag aaacatggtg aaaacaggct gaggttgtca 420
84 gggcagagag ctgaaggtgg ggacagtgc cgcggacccc tctgtgcttg aaagatttcc 480
86 tccacggcct ttgccccagt tgtggggagg tctctgtgca cagcggggaa aatgcttgtg 540
88 tcgcctttgg tgggccatgt cctaattagt ttcctctgct tccctgggaa cttactaagg 600
90 ggcccagagc actgttggaa gtctggttag agtccccaga gagttactct aagttaaaat 660
92 gagccactga ccttggtcct ccttagagga atttcctcga gaacaacaga gataagaaaa 720
94 gaaccggcct ggccaatcct tcaacagctc taga 754
97 <210> SEQ ID NO: 3
98 <211> LENGTH: 1292
99 <212> TYPE: DNA
100 <213> ORGANISM: Homo sapiens
102 <220> FEATURE:
103 <221> NAME/KEY: misc_feature
104 <222> LOCATION: (1244)..(1244) ✓
105 <223> OTHER INFORMATION: n is a,c,g, or t
107 <220> FEATURE:
108 <221> NAME/KEY: misc_feature
109 <222> LOCATION: (1251)..(1251) ✓
110 <223> OTHER INFORMATION: n is a,c,g, or t
112 <400> SEQUENCE: 3
113 cggcacgagg gagatgatga gttgctgagt gtgcacacct ttccggaaca catacacaca 60
115 ccctgctctg ggatcccttg tgaggctgcc ctcatggagt tccccctggc ccagatatgt 120
117 cccaagcac cttccagatc acagacatga cccgcaggag ctgccagaac ctgggctaca 180
119 ctgcggcatc tccccaggcc cgggaggctg cctccagcac agggaatgct gagagggcag 240
121 aggaggtgcc tggagaagga agcctgttcc tgcaggccga gaccgggct tggttccaga 300
123 agaccaggc cactggctc ctgcagcac gggcagcccc tgctggttc catggcttca 360
125 tcaccggag ggaggcagag aggctgctgg agcccaagcc tcaggggtgc tacttggtgc 420
127 ggttcagcga gagcgcggtg accttcgtgc tgacttacag gagccggact tctgcccgc 480
129 acttctctgt ggcccagctc agggacgggc gccacgtggt gctgggcgag gacagcgccc 540
131 acgcgcggct gcaggacctg ctgctgcact acaccgcga cccgctcagc ccctacgggg 600
133 agacgtcac cgagccctc gcccgacaga ctctgagcc tgcaggactt tccctgagga 660
135 ccgaagaatc aaactttgga agcaaaagcc aggacccaaa cccccagtac agcccaatca 720
137 tcaaacaggg gcaagcccca gtcccgatgc agaaagaggg ggccggggag aaggagccct 780
139 ccagctgct caggcccaag cctcccattc ccgccaaacc tcagctgccc ccagaagtct 840
141 acacaatccc tgttccacga caccgcccgg cccacgccc caagccctcc aatcctatct 900
143 acaatgagcc tgatgaaccc atagctttct atgccatggg ccggggcagc cctggggaag 960
145 ccccagcaa catctatgtg gaagtggaa atgagggcct acccgccacc cttgggcacc 1020
147 ctgtcctacg gaagagctgg tccaggcctg tccaggagg ccagaataca ggtggctccc 1080
149 agctgcattc tgagaactct gtgattgggc aaggccctcc cctgccccac cagccccac 1140
151 ccgcctggag acacaccctc cccacaatc tttctagaca ggtgcttcag gacagaggac 1200
W--> 153 aggcattggt tccccctggg cctcctcagt aggcggtaag tctnagcttt ngtgaggata 1260
155 actttggagc atcaggttcc agagacctt ca 1292
158 <210> SEQ ID NO: 4
159 <211> LENGTH: 1323

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/724,676A

DATE: 11/20/2002

TIME: 10:48:50

Input Set : D:\pat.ST25.txt

Output Set: N:\CRF4\11192002\I724676A.raw

```

160 <212> TYPE: DNA
161 <213> ORGANISM: Homo sapiens
163 <220> FEATURE:
164 <221> NAME/KEY: misc_feature
165 <222> LOCATION: (1275)..(1275)
166 <223> OTHER INFORMATION: n is a,c,g, or t
168 <220> FEATURE:
169 <221> NAME/KEY: misc_feature
170 <222> LOCATION: (1282)..(1282)
171 <223> OTHER INFORMATION: n is a,c,g, or t
173 <400> SEQUENCE: 4
174 cggcacgagg gagatgatga gttgctgagt gtgcacacct ttccggaaca catacacaca      60
176 ccctgctctg ggatcccttg tgaggctgcc ctcattggagt tccccctggc ccagatatgt      120
178 ccccaaggga gtcacgaagc ccccatccca accttcagca ccttcagat cacagacatg      180
180 acccgagga gctgccagaa cctgggctac actgcggcat ctccccaggc cccggaggct      240
182 gcctccagca caggggaatgc tgagagggca gaggaggtgc ctggagaagg aagcctgttc      300
184 ctgcaggccg agaccgggc ttggttccag aagaccagg cccactggct cctgcagcac      360
186 ggggcagccc ctgcctggtt ccatggcttc atcaccggg gggaggcaga gaggctgtg      420
188 gagcccaagc ctcaggggtg ctacttggtg cggttcagcg agagcgcgt gaccttcgtg      480
190 ctgacttaca ggagccggac ttgctgccgc cacttcctgc tggcccagct caggacggg      540
192 cgccaagtgg tgctgggcga ggacagcgcc cagcgcggc tgaggacct gctgctgcac      600
194 tacaccgccc acccgctcag cccctacggg gagacgtca ccgagccct cggccgacag      660
196 actcctgagc ctgcaggact ttccctgagg accgaagaat caaactttgg aagcaaaagc      720
198 caggacccaa acccccagta cagcccaatc atcaaacagg ggcaagcccc agtcccgatg      780
200 cagaaagagg gggccgggga gaaggagccc tcccagctgc tcaggcccaa gcctcccatc      840
202 cccgccaaac ctcagctgcc cccagaagtc tacacaatcc ctgttccacg acaccgccc      900
204 gccccacgcc ccaagccctc caatcctatc tacaatgagc ctgatgaacc catagctttc      960
206 tatgccatgg gccggggcag ccctggggaa gccccagca acatctatgt ggaagtggaa      1020
208 gatgagggcc taccgcccac ccttgggcac cctgtctac ggaagagctg gtccaggcct      1080
210 gtcccaggag gccagaatac aggtggctcc cagctgcatt ctgagaactc tgtgattggg      1140
212 caaggccctc ccctgccccca ccagccccc cccgcctgga gacacaccct cccccacaat      1200
214 ctttctagac aggtgcttca ggacagagga caggcatggc ttccccttgg gcctcctcag      1260
W--> 216 tagggcgtaa gtctnagctt tngtgaggat aactttggag catcagggtc cagagacctt      1320
218 tca
219
221 <210> SEQ ID NO: 5
222 <211> LENGTH: 2119
223 <212> TYPE: DNA
224 <213> ORGANISM: Homo sapiens
226 <400> SEQUENCE: 5
227 accgccacca tgaacaagct ttacatcggc aacctcaacg agagcgtgac ccccgcgagc      60
229 ttggagaaag tgtttgcgga gcacaagatc tcctacagcg gccagttctt ggtcaaatcc      120
231 ggctacgcct tcgtggactg cccggacgag cactgggcga tgaaggccat cgaaactttc      180
233 tccgggaaag tagaattaca aggaaaacgc ttagagattg aacattcggg gcccaaaaaa      240
235 caaaggagcc ggaaaattca aatccgaaat attccacccc agctccgatg ggaagtactg      300
237 gacagcctgc tggctcagta tggtagagta gagaactgtg agcaagtga caccgagagt      360
239 gagacggcag tggtagaatg cacctattcc aaccgggagc agaccaggca agccatcatg      420
241 aagctgaatg gccaccagtt ggagaaccat gccctgaagg tctcctacat ccccgatgag      480
243 cagatagcac agggacctga gaatgggcgc cgagggggct ttggctctcg gggtcagccc      540
245 cgccaggctc accttgtggc agcggggggc cagccaagca gcagcaagtg gacatcccc      600

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/724,676A

DATE: 11/20/2002

TIME: 10:48:50

Input Set : D:\pat.ST25.txt

Output Set: N:\CRF4\11192002\I724676A.raw

247	ttcggctcct	ggtgcccacc	cagtatgtgg	gtgccattat	tggcaaggag	ggggccacca	660
249	tccgcaacat	cacaaaacag	acccagtcca	agatagacgt	gcataggaag	gagaacgcag	720
251	gtgcagctga	aaaagccatc	agtgtgcact	ccaccctga	gggctgctcc	tccgcttgta	780
253	agatgatctt	ggagattatg	cataaagagg	ctaaggacac	caaaacggct	gacgaggttc	840
255	ccctgaagat	cctggcccat	aataactttg	tagggcgtct	cattggcaag	gaaggacgga	900
257	acctgaagaa	ggtagagcaa	gataccgaga	caaaaatcac	catctcctcg	ttgcaagacc	960
259	ttacccttta	caaccctgag	aggaccatca	ctgtgaaggg	ggccatcgag	aattgttgca	1020
261	gggccgagca	ggaaataatg	aagaaagtcc	gggaggccta	tgagaatgat	gtggctgcca	1080
263	tgagcctgca	gtctcacctg	atccctggcc	tgaacctggc	tgctgtaggt	cttttcccag	1140
265	cttcatccag	cgcagtcctg	ccgcctccca	gcagcggtac	tggggctgct	ccctatagct	1200
267	cctttatgca	ggctcccag	caggagatgg	tgcaggtgtt	tatccccgcc	caggcagtgg	1260
269	gcgccatcat	cggcaagaag	gggcagcaca	tcaaacagct	ctcccggttt	gccagcgctt	1320
271	ccatcaagat	tgcaccaccc	gaaacacctg	actccaaagt	tcgtatgggt	atcatcactg	1380
273	gaccgccaga	ggcccaattc	aagggtctcc	agtatgatag	ctaagcagaa	ggaaagctga	1440
275	caagctgccc	acttgagcct	tgcccttcac	ctgccttaga	tggctcaggg	aagaatctat	1500
277	ggcaaaactca	aggaggagaa	cttcttttgt	cccaaggagg	aagtgaagct	ggagaccac	1560
279	atacgtgtgc	cagcatcagc	agctggccgg	gtcattggca	aagggtgaaa	aacggtgaac	1620
281	gagttgcaga	atttgacggc	agctgaggtg	gtagtaccaa	gagaccagac	ccctgatgag	1680
283	aacgaccagg	tcacgtgaa	aatcatcgga	catttctatg	ccagtcagat	ggctcaacgg	1740
285	aagatccgag	acatcctggc	ccagggttaag	cagcagaaat	cgagagtgtg	ctctccccgg	1800
287	caggcctgag	aatgagtggg	aatccgggac	acctggggcg	ggctgtagat	caggtttgcc	1860
289	cacttgattg	agaaagatgt	tccagtgagg	aacctgatc	tctcagcccc	aaacaccac	1920
291	ccaattggcc	caacactgtc	tgcccctcgg	ggtgtcagaa	attctagcgc	aaggcacttt	1980
293	taaacgtgga	ttgttttaag	aagctctcca	ggccccacca	agagggtgga	tcacacctca	2040
295	gtgggaagaa	aaataaaatt	tccttcaggt	tttaaaaaca	tgcaagagag	tgttttaatc	2100
297	agccttaaa	gatggttca					2119
300	<210>	SEQ ID NO: 6					
301	<211>	LENGTH: 2040					
302	<212>	TYPE: DNA					
303	<213>	ORGANISM: Homo sapiens					
305	<400>	SEQUENCE: 6					
306	accgccacca	tgaacaagct	ttacatcggc	aacctcaacg	agagcgtgac	ccccgcggac	60
308	ttggagaaag	tgtttgcgga	gcacaagatc	tcctacagcg	gccagtctct	ggtcaaatcc	120
310	ggctacgcct	tcgtggactg	ccgggacgag	cactgggcga	tgaaggccat	cgaaactttc	180
312	tccgggaaag	tagaattaca	aggaaaacgc	ttagagattg	aacattcggg	gccccaaaaa	240
314	caaaggagcc	ggaaaattca	aatccgaaat	attccacccc	agctccgatg	ggaagtactg	300
316	gacagcctgc	tggctcagta	tggtacagta	gagaactgtg	agcaagtga	caccgagagt	360
318	gagacggcag	tggtagaatg	cacctattcc	aaccgggagc	agaccaggca	agccatcatg	420
320	aagctgaatg	gccaccagtt	ggagaacat	gccctgaagg	tctcctacat	ccccgatgag	480
322	cagatagcac	agggacctga	gaatgggcgc	cgagggggct	ttggctctcg	gggtcagccc	540
324	cgccaggctc	accttgtggc	agcggggggc	cagccaagca	gcagcaagtg	gacatcccc	600
326	ttcggctcct	ggtgcccacc	cagtatgtgg	gtgccattat	tggcaaggag	ggggccacca	660
328	tccgcaacat	cacaaaacag	acccagtcca	agatagacgt	gcataggaag	gagaacgcag	720
330	gtgcagctga	aaaagccatc	agtgtgcact	ccaccctga	gggctgctcc	tccgcttgta	780
332	agatgatctt	ggagattatg	cataaagagg	ctaaggacac	caaaacggct	gacgaggttc	840
334	ccctgaagat	cctggcccat	aataactttg	tagggcgtct	cattggcaag	gaaggacgga	900
336	acctgaagaa	ggtagagcaa	gataccgaga	caaaaatcac	catctcctcg	ttgcaagacc	960
338	ttacccttta	caaccctgag	aggaccatca	ctgtgaaggg	ggccatcgag	aattgttgca	1020
340	gggccgagca	ggaaataatg	aagaaagtcc	gggaggccta	tgagaatgat	gtggctgcca	1080

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/724,676A

DATE: 11/20/2002

TIME: 10:48:50

Input Set : D:\pat.ST25.txt

Output Set: N:\CRF4\11192002\I724676A.raw

342	tgagcctgca	gtctcacctg	atccctggcc	tgaacctggc	tgctgtaggt	cttttcccag	1140
344	cttcatccag	cgcagtcctg	ccgcctccca	gcagcggtac	tggggctgct	ccctatagct	1200
346	cctttatgca	ggctcccag	caggagatgg	tgacggtgtt	tatccccgcc	caggcagtgg	1260
348	gcgccatcat	cggcaagaag	gggcagcaca	tcaaacagct	ctccccggtt	gccagcgct	1320
350	ccatcaagat	tgaccacccc	gaaacacctg	actccaaagt	tcgtatgggt	atcatcactg	1380
352	gaccgccaga	ggccaattc	aaggctcagg	gaagaatcta	tggcaaactc	aaggaggaga	1440
354	acttcttttg	tccaaggag	gaagtgaagc	tggagaccca	catacgtgtg	ccagcatcag	1500
356	cagctggccg	ggctattggc	aaaggtggaa	aaacggtgaa	cgagttgcag	aatttgacgg	1560
358	cagctgaggt	ggtagtacca	agagaccaga	cccctgatga	gaacgaccag	gtcatcgtga	1620
360	aatcatcgg	acatttctat	gccagtcaga	tggctcaacg	gaagatccga	gacatcctgg	1680
362	cccaggttaa	gcagcagaaa	tcgagagtgt	gctctccccg	gcaggcctga	gaatgagtgg	1740
364	gaatccggga	cacctgggcc	gggctgtaga	tcaggtttgc	ccacttgatt	gagaaagatg	1800
366	ttccagtgg	gaacctgat	ctctcagccc	caaacaccca	cccaattggc	ccaacactgt	1860
368	ctgcccctcg	gggtgtcaga	aattctagcg	caaggcactt	ttaaactgtg	attgtttaaa	1920
370	gaagctctcc	aggccccacc	aagagggtgg	atcacacctc	agtgggaaga	aaaataaaat	1980
372	ttccttcagg	ttttaaaaac	atgcagagag	gtgttttaat	cagccttaaa	ggatggttca	2040
375	<210> SEQ ID NO: 7						
376	<211> LENGTH: 1990						
377	<212> TYPE: DNA						
378	<213> ORGANISM: Homo sapiens						
380	<400> SEQUENCE: 7						
381	accgccacca	tgaacaagct	ttacatcggc	aacctcaacg	agagcgtgac	ccccgcggac	60
383	ttggagaaa	tggttgccga	gcacaagatc	tcctacagcg	gccagttctt	ggtaaaatcc	120
385	ggctacgct	tcgtggactg	ccgggacgag	cactgggcga	tgaaggccat	cgaaactttc	180
387	tcggggaaa	tagaattaca	aggaaaacgc	ttagagattg	aacattcggt	gccccaaaaa	240
389	caaaggagcc	ggaaaattca	aatccgaaat	attccacccc	agctccgatg	ggaagtactg	300
391	gacagcctgc	tggtcagta	tggtacagta	gagaactgtg	agcaagtga	caccgagagt	360
393	gagacggcag	tggtgaatgt	cacctattcc	aaccgggagc	agaccaggca	agccatcatg	420
395	aagctgaatg	gccaccagtt	ggagaaccat	gccctgaagg	tctcctacat	ccccgatgag	480
397	cagatagcac	agggacctga	gaatgggcgc	cgagggggct	ttggctctcg	gggtcagccc	540
399	cgccaggctc	accttgtggc	agcgggggcc	cagccaagca	gcagcaagtg	gacatcccc	600
401	ttcggtcct	ggtgccacc	cagtatgtgg	gtgccattat	tggcaaggag	ggggccacca	660
403	tcgcaacat	cacaaaacag	accagtcaca	agatagacgt	gcataggaag	gagaacgcag	720
405	gtgcagctga	aaaagccatc	agtgtgcact	ccaccctga	gggtgctcc	tcgcttgta	780
407	agatgatctt	ggagattatg	cataaaggag	ctaaggacac	caaaacggct	gacgaggttc	840
409	ccctgaagat	cctggcccat	aataactttg	tagggcgctc	cattggcaag	gaaggacgga	900
411	acctgaagaa	ggtagagcaa	gataccgaga	caaaaatcac	catctcctcg	ttgcaagacc	960
413	ttacccttta	caaccctgag	aggaccatca	ctgtgaaggg	ggccatcgag	aattgttgca	1020
415	gggccgagca	ggaaataatg	aagaaagtgc	gggaggccta	tgagaatgat	gtggctgcca	1080
417	tgagcctgca	gtctcacctg	atccctggcc	tgaacctggc	tgctgtaggt	cttttcccag	1140
419	cttcatccag	cgcagtcctg	ccgcctccca	gcagcggtac	tggggctgct	ccctatagct	1200
421	cctttatgca	ggctcccag	caggagatgg	tgacggtgtt	tatccccgcc	caggcagtgg	1260
423	gcgccatcat	cggcaagaag	gggcagcaca	tcaaacagct	ctccccggtt	gccagcgct	1320
425	ccatcaagat	tgaccacccc	gaaacacctg	actccaaagt	tcgtatgggt	atcatcactg	1380
427	gaccgccaga	ggccaattc	aaggctcagg	gaagaatcta	tggcaaactc	aaggaggaga	1440
429	acttcttttg	tccaaggag	gaagtgaagc	tggagaccca	catacgtgtg	ccagcatcag	1500
431	cagctggccg	ggctattggc	aaaggtggaa	aaacggtgaa	cgagttgcag	aatttgacgg	1560
433	cagctgaggt	ggtagtacca	agagaccaga	cccctgatga	gaacgaccag	gtcatcgtga	1620
435	aatcatcgg	acatttctat	gccagtcaga	tggctcaacg	gaagatccga	gacatcctgg	1680